REMARKS

Anticipation Rejection of Claims 1-4

In the July 10, 2006 office action, the Examiner rejected claims 1-4 as anticipated by Furman et al., US 6,049,594. Applicants respectfully traverse the rejection.

Claim 1 is directed to an improvement to a voice command platform comprising computer software and a memory comprising a phoneme dictionary and application grammar for an application executing on said voice command platform. The improvement comprises a method which includes steps of obtaining phonemes from audio files comprising spoken names of users of said platform, the spoken names comprising user's speech of their own names; and modifying the phoneme dictionary and application grammar based on the phonemes obtained from the audio files. Applicants have amended claim 1 to recite that the improvement is a method in order to more clearly have the claim fall into one of the statutory classes of subject matter, and to more clearly recite that the audio files comprising spoken names of users is the users' speech of their own names.

As noted in the specification, e.g., at pages 3-5 and 20-21, this invention addresses a need to have a voice command platform provide a facility which better recognizes spoken names, to take in to account the myriad pronunciations of users' names due to ethnic, age, dialect, accent, and other linguistical variations in names of a population of users. This is achieved by obtaining audio files comprising speech input comprising the users' speech of their own names (typically in a training session in response to prompts) and modifying the phoneme dictionary and application grammar based on phonemes obtained from the audio files.

Furman et al. does not anticipate because it does not teach modification of a phoneme dictionary based on the spoken names of the users, i.e., the user's speech of their own name. Furman describes a training process where the user is prompted to speak the name *of the persons that they frequently call* (Example 3; Figure 8, col. 10 lines 10 et seq.), *not their own name for use when people call them*.

The description of Figure 3, step 70 cited by the Examiner, recites to a first embodiment in which there is no training at all – the "speech training processor" constructs phoneme strings for names based on a database of phonemes and translation of text to phonemes (not on the basis of speech input), moreover the names are names of persons the user frequently calls, not the spoken name of the user. See col. 5 lines 35-45 ("Next, processor 5 identifies the name of the called party for each of the 20 most frequently called numbers . . . The name of each called party is retrieved by processor 5 as a text string . . ."). Furman does not disclose a training or other process in which the user's speech of their own name is obtained, and the phoneme dictionary and application grammar modified accordingly.

Since claim 1 is clearly not anticipated, the rejection of dependent claims 2-3 should also be withdrawn.

Claim 4 is a method claim which recites:

. . .

[a] prompting a user to provide speech input comprising their spoken name;
[b] receiving said speech input and saving said speech input as an audio file;
[c] converting said speech input to a set of phonemes;
[d] modifying said application grammar based on said set of phonemes; and
[e] modifying said phoneme dictionary based on said set of phonemes.

(brackets added for ease of explanation).

Clearly, step [a] refers to providing speech input comprising the user's own spoken name, not the name of some third party. The Examiner cites to col. 9 line 11 of Furman for step [a]. This passage recites that the system "prompt[s] the customer to speak a name which the customer would like to say when voice dialing a specific person." The text at col. 9 line 19-25 clearly indicates that the training process is describing speech of the name of the party they are calling, *not their own name.* In this respect, the reference is describing a second embodiment in which the system acquires speech of the name of the party they are calling, whereas the embodiment of Figure 3 and cols. 5-6 describes a process in which the name is ascertained without any speech input at all. However, in both the passage at col. 9 and elsewhere in the document, the reference is referring to acquiring speech input for the calling party, not the user's pronunciation of their own name. Accordingly, Furman does not anticipate claim 4.

Obviousness Rejection of claims 5 and 6

The Examiner rejects claim 5 and 6 as obvious over Furman et al. in combination with Curt et al., US 6,438,520. Claim 5 recites, among other things, a step of conducting a tutorial process, said tutorial process prompting a user to provide speech input comprising their spoken name and receiving said speech input and saving said speech input as an audio file.

In the rejection of claim 5, the Examiner repeats the error of analysis of the reference and the claim as in claims 1 and 4. In particular, the Examiner cites to col. 14 lines 31-36 for a teaching of prompting the user to provide speech input of their own spoken name. The passage at col. 14 cited the Examiner deals with how to provide

name information *for frequently called numbers,* such as unlisted numbers. Unlisted numbers pose a special problem for Furman's Figure 3 embodiment because there is no text file of the name of the called party to use to convert to one or more phonemes. The passage at col. 14 suggests that the user could be prompted to "speak a name (label) which is used to train an HMM (hidden Markov model) or select a sequence of phonemes using automatic speech recognition techniques." The passage also states that such a procedure could be used for all frequently dialed numbers. Thus, Furman is again teaching at most providing speech input of *the name of the called party*, n*ot their own spoken name*.

Curt et al. is directed to a method for recognizing names of parties that leave voice messages, and teaches away from a training process in which the user of the system is prompted to speak their own name and responsively modifying a phoneme dictionary and application grammar. See passage at col. 5 cited by the Examiner, example of speech of name "Rafid" (" . . . the various embodiments of the present invention will also automatically dial the number associate with "Rafid", as spoken by the subscriber. In these instances, *neither the incoming caller nor the subscriber has trained the name "Rafid" as stored in the message list*." col. 5 lines 22 -27 (emphasis added)). Rather, Curt describes performing a phonetic transcription of the speaker's name, created using a speaker-independent, hidden Markov model (HMM) having an unconstrained grammar in which any phoneme may follow any other phoneme. (col. 5 lines 35-45). The speech is utilized to select a closest match to an existing phoneme pattern, if any, using the speaker independent HMM based model. Based on likelihood of fit parameters, the disclosure

determines whether the incoming speech pattern matches (collides) with any existing phoneme pattern.

Accordingly, applicants submit that the combination of Furman et al. and Curt et al. fails to render obvious the subject matter of claim 5. Claim 6 is also allowable by virtue of claim dependency.

Favorable reconsideration of the application is requested.

Respectfully submitted.

McDonnell Boehnen Hulbert & Berghoff LLP

Date: 8/21/06 By: _____

Thomas A. Fairhall
Reg. No. 34591

CERTIFICATE OF MAILING

The undersigned hereby certifies that the foregoing RESPONSE TO OFFICE ACTION MAILED JULY 10, 2006, is being deposited as first class mail, postage prepaid, in an envelope addressed to Mail Stop Amendment, Commissioner for Patents, P.O. BOX 1450 Alexandria VA 22313-1450 on this 21st day of August, 2006.

_____

Thomas A. Fairhall